

The Effect of Training on Fricative Production in Second Language
Speakers

A Senior Honors Thesis

Presented in Partial Fulfillment of the Requirements for graduation with research
distinction in Speech and Hearing Sciences in the undergraduate colleges of The Ohio
State University.

By

Amy E. Loscko

The Ohio State University

May 2009

Project Advisors: Dr. Robert Fox and Dr. Ewa Jacewicz

TABLE OF CONTENTS

INDEX TO FIGURES	3
ACKNOWLEDGEMENTS	4
ABSTRACT	5
CHAPTER 1. Introduction and Literature Review	6
CHAPTER 2. Methodology	11
CHAPTER 3. Results	14
CHAPTER 4. Conclusion and Discussion	20
Appendix	22
REFERENCES	23

INDEX TO FIGURES

Figure 3.1	16
Figure 3.2	17
Figure 3.3	18
Figure 3.4	19

Acknowledgments

I would like to thank and acknowledge Dr. Robert Allen Fox and Dr. Ewa Jacewicz for their unending patience, guidance, and support with this project. I would also like to thank John Grinstead for sparking my interest in research and for serving on my defense committee. Finally, I am forever grateful to my family and friends who have supported me during this challenging year.

This project was financially supported by the Department of Speech and Hearing Science at The Ohio State University.

Abstract

Speech produced by L2 (second language) speakers of American English can differ greatly from native speakers, largely due to the interaction of the phonetic characteristics of their native language and those of English in the L2 acquisition process (Flege, 1995). It is well known that native speakers of Korean have difficulties acquiring the English /s/-/ʃ/ fricative contrast, which depends largely on the position of the fricative in the word. In English, you can have both fricatives contrast in almost any phonetic context, but this is not the case in Korean. Previous research has examined the nature of this acoustic contrast in English with native adults and children; however, no such acoustic measures have been examined for L2 speakers. In addition, we do not know if short-term “training” on the fricative contrast will improve L2 productions (i.e., make the acoustic characteristics of these fricatives more similar to native English productions). This research project examines how well native Korean speakers improved their production of English fricatives after two training sessions which provided explicit instruction as to proper pronunciation of the fricatives in different positions in the word. Fricative productions by native English speakers served as a control. Results suggest that L2 speakers can significantly improve their pronunciation of English fricatives with very short periods of training.

Chapter 1

Introduction and Literature Review

This paper reports data that is an extension of an ongoing project being conducted in the Speech & Acoustics Labs at The Ohio State University. The focus of the research is the effect of training on fricative production in second language speakers. Specifically, it will examine short-term learning effects on the acoustic characteristics of the /s/ and /ʃ/ fricatives in word-initial, word-medial, and word-final positions in real English words. The acoustic measures used in the study will determine whether the productions of these fricatives become more like native productions following a short period of training.

Speech production research aims to define acoustic properties of speech sounds that occur in a natural language. Previous research has been conducted on native English adults and children to examine the acoustic characteristics of English fricatives, but no such in-depth research has been conducted on speakers who are second language learners of English (L2). For children, it was found that young children's sibilant fricative articulations are not differentiated in an adult-like manner, but are continuing to be "fine-tuned" as the child matures (Nissen and Fox, 2005). This poses the question as to how effective short-term training is on the improvement of fricative productions in L2 speakers.

Fricatives are consonants produced with a very narrow constriction in the oral cavity. As cited in Jongman et al. (2000), a rapid flow of air through the constriction (the position of which depends on the particular fricative) creates turbulence in the flow, and the random velocity fluctuations in the flow act as a source of sound (e.g., Stevens, 1971,

1998; Shadle, 1990). Generally, English fricatives are categorized into four classes according to their place of articulation: labiodental /f/, /v/, inter-dental /θ/, /ð/, alveolar /s, z/, and palato-alveolar /ʃ/, /ʒ/.

The present study will examine the acoustic contrast of the alveolar /s/ and palato-alveolar /ʃ/. It is well known that native speakers of Korean have difficulties acquiring the English /s/, /ʃ/ fricative contrast. Unlike native English, the Korean language lacks the /ʃ/ phoneme. This linguistic difference between English and Korean will interfere with the Korean speakers' ability to contrast the two English fricatives in both perception and production. The difficulties in pronunciation depend largely on the position of the fricatives in the word.

General principles of phonology predict that learners will be able to acquire a contrast in some positions before they can acquire the contrast in other positions. It is expected that Korean learners will acquire a contrast in word-initial positions before they acquire the contrast in the morphologically derived word-medial position and word-final position. Word-initial positions in English are syllable-onsets and the morphologically derived word-medial position and word-final position are both syllable codas. Contrasts in onsets have been found to be acquired before contrasts in codas (Eckman and Iverson 1994).

General principles of phonology also make predictions about L2 learners splitting native allophones into two L1 differentiated phonemes. The notion is that learners will be able to suppress the use of the allophonic rule across a morpheme boundary only if they are also able to prevent the rule from applying within morphemes (Eckman, Elreyes and Iverson 2003). In other words, an L2 learner will acquire the phonemic contrast in

morphologically-derived environments only if that learner has acquired the contrast in basic environments, such as word-initial positions.

Speech produced by L2 speakers of American English can differ greatly from native speakers, largely due to the interaction of their native phonetic categories in the L2 acquisition process (Flege, 1995). The Speech Learning Model developed by Flege posits that L2 learners are prevented from creating a separate category for L2 sounds which are similar to L1 sounds. The absence of a separate category is a result of the similarities from the native L1, thus the two categories are “merged” under one (Flege, 2003). In this manner, the inability of L2 learners to produce authentic L2 phonemes is explained. As a result, non-native accented speech can be difficult to understand by native English speakers.

It is not well understood how much L2 speakers can improve their production accuracy of non-native sounds as a result of training and exposure. Increased acoustic variability can cause some difficulty in identifying sounds as the precise acoustic characteristics of a speech segment may change as a function of phonetic context (e.g., Mullennix et al., 1989). For example, the contrast between /s/ and /ʃ/ may be reduced before a rounded vowel (in this context the /s/ will be rounded and it will lower its spectral mean—the /ʃ/ is usually rounded in all contexts). L2 learners may acquire segmental contrasts better and retain them longer when they are exposed to a variety of contexts, including different positions of the sounds in a word and varied morphological complexity of a word.

This study aims foremost to determine whether there is a significant difference between the two voiceless fricatives of interests: alveolar /s/ and palato-alveolar /ʃ/

fricative productions in English words by Korean speakers compared with English speakers and whether this difference will improve in the Korean speaker group following training. Specifically, this research will determine if the fricative productions of the L2 speakers can become more English native-like following a short period of training. Of similar interest was the examination of learning effects across different positions of the segmental contrast in a word by comparing fricative characteristics in word-initial, and word-medial positions.

Previous studies have examined which spectral variables are successful in differentiating the acoustic properties of English fricatives. The present study will use spectral and temporal measures including spectral mean, spectral skewness, peak LPC and fricative noise duration to analyze the contrast of English /s/ and /ʃ/. These acoustic variables were successfully used in previous research with adults and children (Jongman et al., 2000; Nissen and Fox, 2005; Nittrouer, 1995), but not with L2 speakers. The acoustic measures used in the study will also determine whether the predictions about the order of acquisition are true (in terms of derived vs. non-derived morphological contexts) and how learners' performance improves with training across differing positions in the word.

It is expected that following a short-term training period, native Korean speakers will be successful in improving the classification rate of the alveolar /s/ and palato-alveolar /ʃ/ English fricatives. The present L2 speakers are expected to perform on the fricative contrast in a manner similar to children's production, so that their articulations are not differentiated in a native-like manner, but will continue to be "fine-tuned" as the

L2 learners gain more experience with English. Improvement is expected to be greatest in word-initial positions.

Methodology

A. Stimulus materials

The speech samples from the Korean speakers were recorded for a previous study in the Speech Perception & Acoustics Labs at The Ohio State University. A set of 60 target words were recorded in total. For the purposes of this study, 12 target words, all of which contained a voiceless fricative of interest, were used to examine fricative acoustic properties. All words were existing lexical items in English. Each target word contained /s/ or /ʃ/ in two different positions in a morphologically basic word: initial (e.g., *see/she*) and final (e.g., *bus/rush*) and in one additional position (medial, at the juncture with another morpheme) in morphologically derived words such as *crossing/washing*. In the word-initial position, /s/ or /ʃ/ occurred before a high front vowel which was also the case for the derived words, in which it was followed by the suffix *-ing*. [The barred-i in “ing” is fairly high and front.] A complete list of the target words used in the production task can be found in Appendix A.

B. Speakers

Ten native speakers of Korean served as participants of the study. All were students at the City University of Incheon, Korea, ranging in age from 19-25 years of age. There were seven male speakers and three female speakers. All participants learned English in a formal school setting (for 8-10 years) and none of them had any direct informal experience with English. The testing and collecting of the recordings were done in two sessions. In the first session, each subject completed a production task to elicit the

pre-training data. In the second session, subject was trained on the production of the /s/-/ʃ/ contrast. The training was administered to the subject in a way similar to the production task, using appropriate pictures and verbal models which were repeated and learned by the subject. After the training, each subject completed a posttest production task which was the same as the pre-training task. For the purposes of this research, only male data will be discussed.

Ten native English speakers elicited a production task for a basis of control. All participants were living in Wisconsin at the time of recording and were native speakers of English. Participants ranged in age from 19-42 years of age. There were six male subjects and four female subjects. For the purposes of this research, only male data will be discussed.

All recordings were made using a Shure SM10A head-mounted microphone positioned approximately 1-inch from the speaker's lips. All recordings were done using a 44.1 kHz sampling rate with 16-bit quantization. All speech samples were recorded directly to the computer's disk drive under the control of a specially written Matlab program.

C. Acoustic analysis of pretest and posttest productions

Measurements of fricative onset and offset were determined using waveform display along with a spectrogram analysis. The segmentation values were inputted into a text file (in ms). A MATLAB program, created by one of the thesis advisors, was used to display the segmentation marks superimposed over a display of the token's waveform. This display was used to check if all measurements are correct. Corrections were made

to the measurements, if necessary. The following acoustic measures were then computed using a separate MATLAB program: Fricative duration, normalized amplitude, and a set of spectral measures including spectral slope, spectral mean, spectral variance, spectral skewness, spectral kurtosis and peak LPC. This set of measures (shown by Jongman, et al., 2000, Fox & Nissen, 2005 and Nissen & Fox, 2005 to accurately discriminate among voiceless fricatives in American English) will be collected for both pretest and posttest productions. For the purposes of this study, only spectral mean, spectral skewness, peak LPC, and fricative duration will be discussed.

SPSS statistics version 17.0 was used to determine two sets of statistical analysis. For each of the variables of interest, two sets of statistics were computed. First the data from both the Korean speakers (pre-test productions) and the English speakers were compared using a 3-way ANOVA with the within-subject factors *segment* and *position* and the between-subject factor *language*. If significant language differences are found, the differences in variable values between /s/ and /ʃ/ in pre- and post-training productions from the Korean speakers were analyzed using a 2-way ANOVA with the within-subject factors *segment* and *position* in order to determine if the short training session improved the L2 speakers' fricative productions.

Results

For each of the variables of interest, two sets of statistics were computed. First the data from both the Korean speakers (pre-test productions) and the English speakers were compared. If significant language differences are found, the differences in variable values between /s/ and /ʃ/ in pre- and post-training productions from the Korean speakers were analyzed. Only significant effects are discussed.

1. *Spectral Mean*

Overall, the spectral mean of /s/ was significantly higher than that of /ʃ/ ($p < .001$); of greater interest is the fact that there was a significant segment by language interaction ($p = .001$)—the spectral means for /s/ and /ʃ/ for the Korean speakers were not as separated as for the English speakers. However, the spectral separation was increased significantly following training ($p = .034$). There was no significant effect of position.

2. *LPC Peak*

Overall, the LPC peak of the fricatives was significantly higher for /s/ than for /ʃ/ ($p < .001$) but again there was a significant segment by language interaction ($p = .001$) as the difference in peak frequency was much smaller for Korean speakers than for English speakers. The /s/- /ʃ/ difference in LPC peak was significantly increased in the post-training condition ($p = 0.30$). Again, there was no significant effect of position.

3. *Spectral Skewness*

Overall, the spectral skewness for /s/ was significantly smaller than for /ʃ/ ($p=.001$).

Again, there was a significant segment by language interaction ($p=.005$). However, Korean speakers did not show a significant improvement in spectral skewness following training.

4. *Duration*

The duration of the fricatives, overall, was longer in initial position than in medial position ($p=.001$); however, the significant position by language interaction ($p=.002$) showed that this was true primarily for English speakers, not Korean speakers (whose mean durations were the same in both position). Korean speakers did not significantly change their fricative durations to match the English pattern following training (n.s.).

Spectral Mean of /s/ and /ʃ/ fricative productions by native Korean and native English speakers

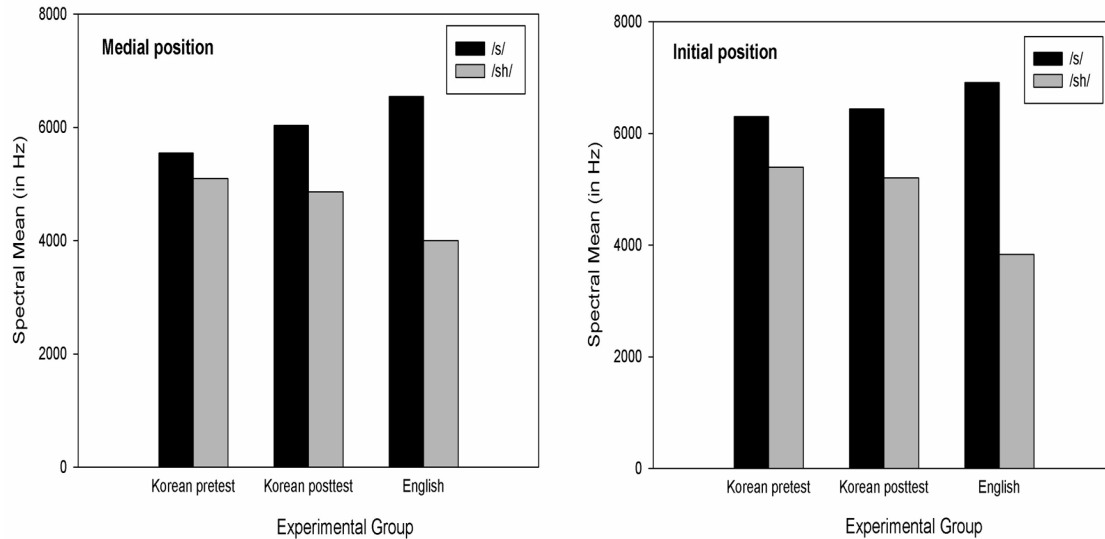


Figure 3.1. Spectral Mean (in Hz) as a function of speaker group and segment: non-palatal (/s/) and palatal (/ʃ/).

As can be seen, there are clear differences in the spectral mean for Korean native speakers compared to English native speakers. Of particular interest is the proportion of segmental differences for the segment by language interaction. For Korean speakers, pretests show a more similar (/s/) and (/ʃ/) spectral mean in both initial and medial word positions. Posttests show a change in both segments such that the spectral mean is closer to that of English native speakers. This change in production of non-palatal (/s/) and palatal (/ʃ/) before and after shows the significant effect training has on fricative production.

Peak LPC of /s/ and /ʃ/ fricative productions by native Korean and native English speakers

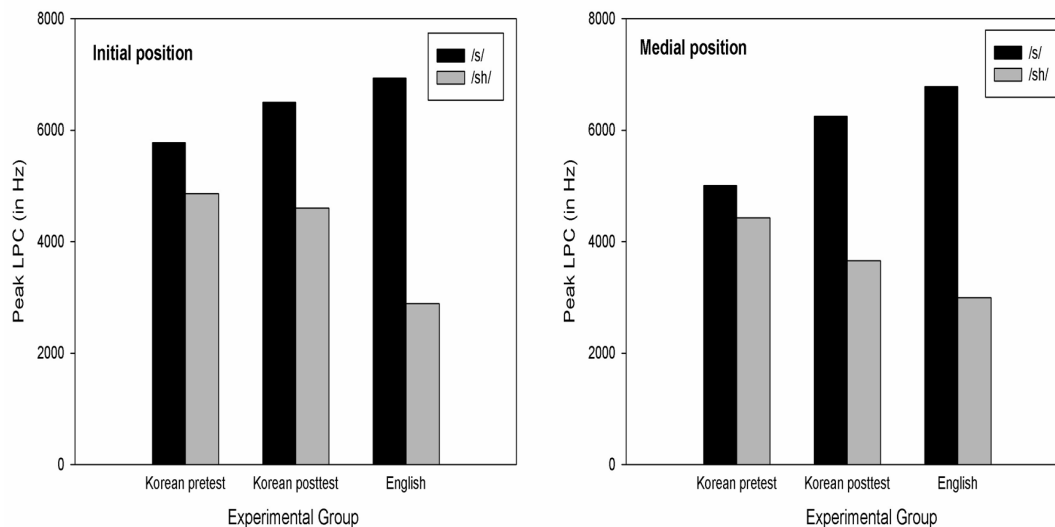


Figure 3.2. Peak LPC as a function of speaker group and segment: non-palatal (/s/) and palatal (/ʃ/).

As can be seen, there are clear differences between Peak LPC of segments /s/ and /ʃ/. Overall, peak LPC for /s/ is much higher in both initial and medial word positions. It is especially important to note the segment by language interaction. Improvement in peak LPC can be seen as fricative production improvements because of the increasing difference between non-palatal /s/ and palatal /ʃ/, typical of native English.

Spectral Skewness of /s/ and /ʃ/ fricative productions in native Korean and native English Speakers

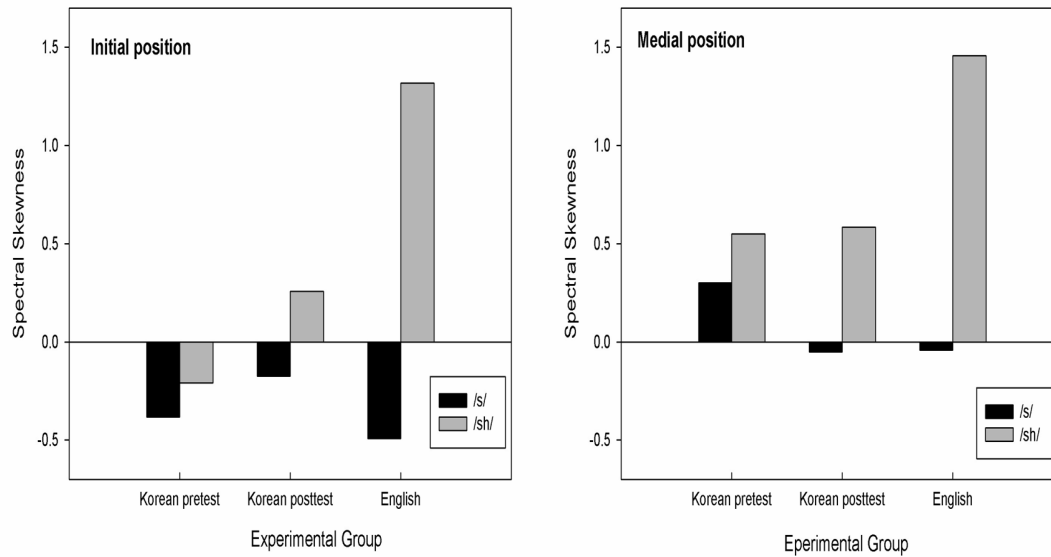


Figure 3.3. Spectral skewness as a function of speaker group and segment: non-palatal (/s/) and palatal (/ʃ/).

These figures show a very different pattern in the spectral skewness of (/s/) and (/ʃ/) for the two speaker groups. In particular, Korean speakers had much smaller values for spectral skewness for / / than did the native English speakers. There seemed to be a change toward the native values for /ʃ/ following training, but not for /s/.

Fricative duration for /s/ and /ʃ/ by native Korean and native English speakers

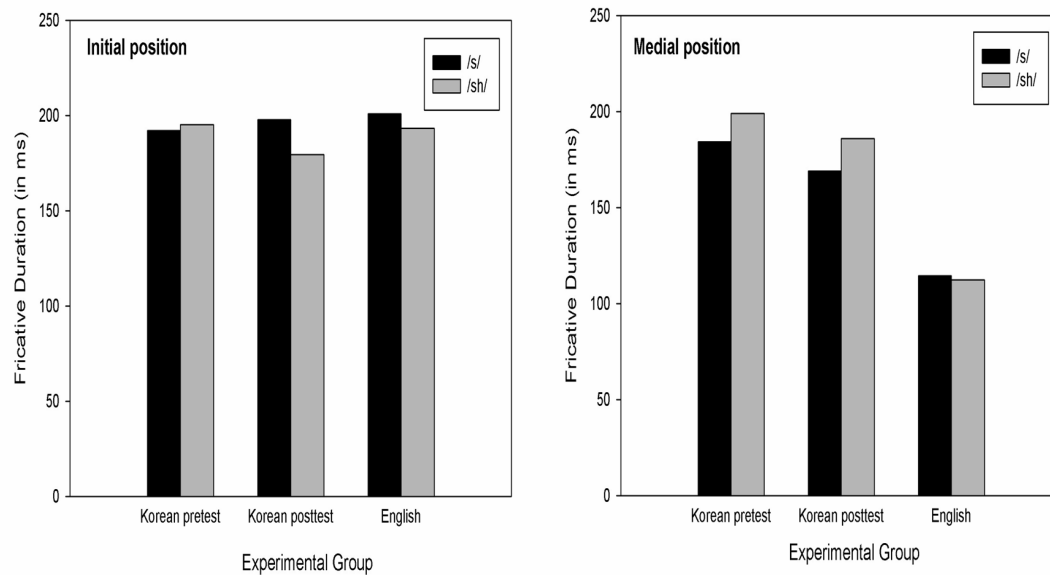


Figure 3.4. Fricative duration as a function of speaker group and segment: non-palatal (/s/) and palatal (/ʃ/).

These figures show an overall tendency for fricative duration to be longer in initial positions than in medial positions. One is tempted to suggest that there is a tendency for the Korean fricative durations to become somewhat more native-like following training, but it was found to be not statistically significant.

Chapter 4

Conclusions and Discussion

These data demonstrate that Korean speakers can make significant improvement to their productions of the voiceless English fricatives /s/ and /ʃ/ following a short training period. In particular, the spectral means and LPC peaks (a primary acoustic difference between these two fricatives) became more native-like (i.e., the values for /s/ increased and for /ʃ/ decreased). There was a slight tendency for the values of spectral skewness and fricative duration for the Korean posttest to change towards the English control, but were found not to be statistically significant. However, we failed to find a predicted difference between initial and medial positions for any of the acoustic measures examined. One might, therefore, infer that the acoustics cues and position within the word are not highly distinctive for these two phonemes. However, it is possible that the classification of these phonemes could improve with alternative spectral parameters. Additionally, our analysis focused on the measurements made at the midpoint of the fricative, which may have been inadequate in observing all relative spectral differences.

As indicated, these results are preliminary in that the speaker sample and word list were rather limited. Further research could prove to be significant for the improvement in the classification of these phonemes by spectral variables if speech samples were increased. Other questions raised by the findings of this study could include what is the minimal duration of training period at which one would see an improvement in L2 fricative productions and also what is the relationship between the effect of training on male and female L2 speakers. Further investigation into the effect of training on fricative

production could also be beneficial if the interaction of L2 speech on one's native language was examined.

While the results of this research provide interesting insight into the effect of training on speech productions by L2 speakers, further studies are necessary to better understand the interaction of native and L2 languages. The results of this study may have implications for the field of research in speech production, how one understands the effects of training on production, as well as ultimately leading to improved strategies and tools for language teaching.

Appendix

The following twelve words were used as target words for both native Korean pretest, posttest as well as for the native English speakers during the production tasks. Each speaker recorded the following tokens with /s/ or /ʃ/ in three different positions within the word. The current study only examines fricatives in Initial and Medial positions.

Initial

1. See
2. She
3. Seat
4. Sheet

Medial

1. Crossing
2. Washing
3. Busing
4. Rushing

Final

1. Cross
2. Wash
3. Bus
4. Rush

References

- Eckman, F. & G. Iverson. 1994. Pronunciation difficulties in ESL: Coda consonants in English interlanguage. In M. Yavas (ed.) *First and second language phonology*, 251 – 265.
- Eckman, F., A. Elreyes, & G. Iverson. 2003. Some principles of second language phonology. *Second Language Research* 19, 169-208.
- Flege, J. E. 1995. Second language speech learning: Theory, findings and problems. In W. Strange (Ed.), *Speech production and linguistic experience: Issues in cross-language research*, 233-277.
- Forrest, K., Weismer, G., Milenkovic, P., & Dougall R. (1988). Statistical analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America* 84, 115-123.
- Flege, J., Schirru, C. & MacKay, I. (2003). Interaction between the native and second language phonetic subsystems. *Speech Comm* 40, 467-491.
- Fox, R. A., and Nissen, S. L. (2005). Sex-related acoustic changes in voiceless English fricatives. *Journal of Speech, Language & Hearing Research* 48, 753-765.
- Jongman, A., Wayland, R., and Wong, S. (2000). Acoustic characteristics of English fricatives. *Journal of the Acoustical Society of America* 108, 1252-1263.

Nissen, S., and Fox, R. A. (2005). Acoustic and spectral characteristics of young children's fricative productions: A developmental perspective. *Journal of the Acoustical Society of America* 118, 2570-2578.

Shadle, C. H. (1990). Articulatory-acoustic relationships in fricative consonants. *Speech Production and Speech Modeling*, edited by W. Hardcastle and A. Marchal (Kluwer, Dordrecht, 1990), 187–209.

Stevens, K. N. (1971), "Airflow and turbulence noise for fricative and stop consonants: Static considerations", *J. Acoust. Soc. Am.* 50, 1180-1192.

Stevens, K.N. (1998), *Acoustic Phonetics*, Cambridge MA: MIT Press, Chapter 1.